On the Performance of Digital Modulation Systems That Expand Bandwidth

By V. K. PRABHU

(Manuscript received December 19, 1969)

It is well known that protection against additive gaussian noise can be obtained in m-ary digital modulation systems by expanding bandwidth or by increasing the channel signal-to-noise ratio. It is also well known that arbitrarily small error probabilities can be attained in digital systems by using long and complex encoding and decoding procedures. Based on the results of Shannon and Slepian, we derive for an optimal system lower bounds to the channel signal-to-noise ratio for various probabilities of error, for various bandwidth expansions, and for a processing interval not greater than the signaling interval of the source. It is assumed that all m characters have equal a priori probabilities and that maximum likelihood detection is used in the receiver. For a bandwidth expansion of two, and for equal energy code words, we also show that the performance of a coherent phase-shift keyed system is as good as that of the optimal system.

I. INTRODUCTION

Various m-ary digital modulation schemes [such as coherent phase-shift keying (CPSK), differentially coherent phase-shift keying (DCPSK), frequency-shift keying (FSK), and others] are currently under investigation for use in satellite, terrestrial, and other radio communication systems.[1,2,3] In such systems, the transmission channel is noisy, and bandlimited, and one is interested in finding an optimum form of modulation for the transmission of information from one point to another. By optimum form of modulation, we mean that we would like to transmit (with a given error rate) as much information as possible in a given band of frequencies and for a given amount of (channel) signal power.

The complexity of the equipment required for particular kinds of modulation, or other considerations in the system, may rule out these optimum transmission schemes in favor of simple and suboptimum

schemes of modulation and demodulation. In order to compare the performance of these simpler practical modulation systems with that of the optimal systems, it is first essential to investigate the performance of these optimal systems.

In this paper we shall assume that the noise in the channel available for communication is gaussian and has a uniform power spectral density over all useful frequency bands. [In terrestrial systems, in the frequency bands above 10 GHz, close spacings of the repeaters are almost always mandatory for reliable communication (during fading conditions produced by rain).[2] If low noise receivers are used in the system, it is possible that the total interference power (due to co-channel and adjacent channel interferers) received by the system may be very much larger than the (thermal) noise power in the system.[3] In this case, note that the total noise corrupting the channel may not be assumed gaussian in all modulation systems (especially if the number of interferers is small).[4,5]]

It is well known that protection against (additive) white gaussian noise can be obtained in m-ary digital modulation systems by expanding bandwidth and/or by increasing the channel signal-to-noise (power) ratio. In fact, Shannon has shown that it is possible to transmit with arbitrarily small error probability the output of a discrete source of entropy R over a channel of bandwidth W perturbed by additive white gaussian noise of average power N by using signals of average power S provided R is less than $W \log_2 (1 + S/N)$ b/s.[6,7] However, such a transmission scheme may involve long and complex encoding and decoding procedures, and to attain these low error rates it may be necessary to provide large storage (or long delay) in the transmitting and receiving equipment. Practical modulation systems presently used for large scale communication do not in general have such large storage capabilities or unlimited bandwidths. Also, we must note that there are practical limitations on the average power of a transmitter and the power that can be received by a receiver.

Since most of the practical modulation systems have a certain bandwidth expansion n and since bandwidth expansion usually improves the (noise) performance of the system, we shall now investigate the optimum performance of digital modulation systems that have a (channel) bandwidth expansion n, a finite channel signal-to-noise power ratio $S/N = \rho^2$, and a processing interval which cannot exceed one signaling interval T of the source.* The message source is assumed to

* It is assumed that the time interval over which the channel is used in decoding one message symbol cannot exceed T.

be of bandwidth W_0, and the channel bandwidth is assumed to be W. Further, it is assumed that all m characters (the output of the discrete message source consists of m different characters or symbols) have equal *a priori* probabilities and that the characters generated by the message source are statistically independent of each other. We also assume that a maximum likelihood detection scheme is used in the receiver.

For such a system, we shall evaluate the lower bounds to the character error probability $P(m, n, \rho^2)$ of the optimal system so that we can compare its performance with that of any practical modulation system. For a given error rate, the difference between the signal-to-noise ratio required by the optimal system and that required by the practical system will then be a measure of the quality of performance of the practical modulation system.

Here, we would like to note that our approach is identical to that of Slepian in Ref. 8 in which he evaluated upper and lower bounds to $P(m, n, \rho^2)$ for n odd, $n \geqq 5$, and $m = 128$ (numerical results for $m = 32$, 64, and 256 can also be found in an unpublished memorandum by Slepian[9]) when $1 \geqq P(m, n, \rho^2) \geqq 10^{-5}$.* In addition to giving numerical results when $P(m, n, \rho^2) < 10^{-5}$ (error rates as low as 10^{-9} are desired in some digital modulation systems[2]) for $m = 2, 4, 8, 16$ and 32, we give a method of evaluating upper and lower bounds to $P(m, n, \rho^2)$ for all values of $n \geqq 2$. We also point out the special significance of $n = 2$, and give closed form solutions for the lower bound when $n = 2$, 3 and 5.

II. COMMUNICATION SYSTEM MODEL†

The m-ary digital modulation system that we shall consider in this paper is assumed to have a signaling interval T. Since we assume Nyquist rate signaling, we shall assume that

$$T = \frac{1}{2W_0}, \tag{1}$$

where W_0 is the bandwidth of the message source. Every T seconds, the message source generates one of m characters or symbols. Since the characters generated by the message source are assumed to be statistically independent, the entropy R of the message source is given by

$$R = 2W_0 \log_2 m \text{ b/s}. \tag{2}$$

* Note that Slepian uses $P(m, n, \rho^2)$ in determining the threshold in *analog* modulation systems that expand bandwidth.

† Compare the communication system model that we discuss in this paper to that given in Ref. 8.

If S is the average power in the channel of bandwidth W, it follows from Ref. 6 that it is possible to transmit with arbitrarily small error probability if and only if

$$2W_0 \log_2 m \leqq W \log_2 (1 + S/N) \tag{3}$$

when there are no restrictions on the way in which the transmitter and receiver operate.

Equation (3) can be shown to yield

$$S/N \geqq m^{2/n} - 1 \tag{4}$$

where

$$n = \frac{W}{W_0} = 2TW = \text{the bandwidth expansion factor.} \tag{5}$$

The lower bound to S/N given by equation (4) is then the smallest signal-to-noise ratio required to transmit (with arbitrarily small error probability) an m-ary digital signal through a channel of bandwidth expansion n when there are no restrictions on the transmitting and receiving equipments. This lower bound to the signal-to-noise ratio S/N is shown in Fig. 1.

In the communication system model we are considering in this paper, there is no provision for the storage of a large number of characters, and hence it is to be expected that we will need signal-to-noise ratios much larger than those given in Fig. 1 (when arbitrarily low error rates are desired). Since the processing interval of our communication system is assumed not to exceed one Nyquist interval (corresponding to the message source), the channel signal corresponding to time T can be used to decode one and only one message symbol.* If the bandwidth of the channel is W, the channel signal can be completely specified by samples taken every $1/2W$ seconds. In the Nyquist interval T, there are then $n = 2WT$ channel samples.† We are then assuming that n channel samples are used to decode one message character, or that each of the m message symbols are mapped into a channel vector having n components.‡ (In the error-free transmission scheme of Shannon, ℓ, $\ell \geqq 1$, successive message symbols are mapped into one channel vector. By making ℓ sufficiently large, by choosing the m^{ℓ} channel vectors appropriately and by decoding appropriately at the receiver, the results given

* This is equivalent to saying loosely that the communication system does not have storage capability for more than one Nyquist interval.[8]

† There are certain subtle points involved in this assumption. Some of these points and their implications are discussed in Refs. 10 and 11.

‡ That is, we construct a dictionary that associates with each of the m message symbols a particular n-dimensional channel vector.

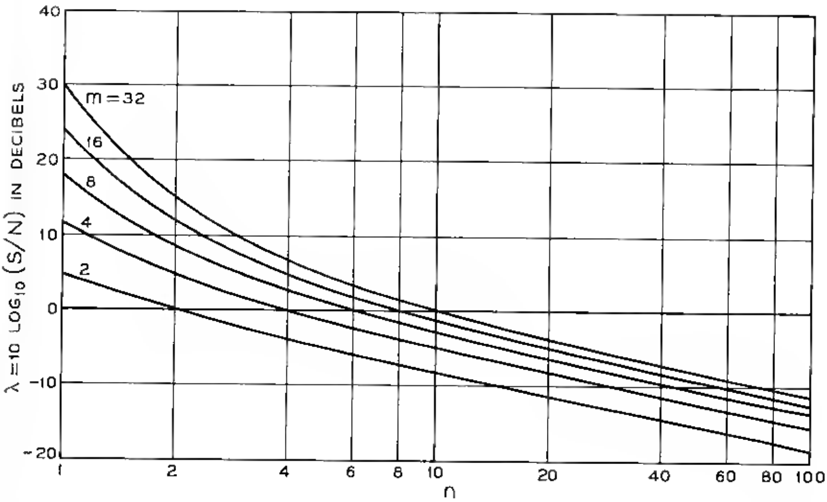

Fig. 1—Lower bound to the signal-to-noise ratio for different bandwidth expansions and ideal signaling.

by Shannon may be obtained.[8] It is to be noted that it is essential to store ℓ message characters before we can generate the appropriate channel vector. In our scheme of transmission, we put $\ell = 1$, and investigate the optimum performance of the system.)

As far as the average power in the channel is concerned, the channel vectors can be chosen in a variety of ways.[12] Since all the characters are assumed to have equal *a priori* probabilities, we will make the assumption that all of them have the same average power S (or the same energy ST).*

Since the noise corrupting the channel is assumed to be white, each component of each channel vector is perturbed independently by the addition of a gaussian variate of mean zero and variance N.

III. EVALUATION OF PROBABILITY OF ERROR $P(m, n, \rho^2)$

Since the channel vectors corresponding to different message symbols have the same average power S, all these n-dimensional vectors ter-

* Other types of restrictions (such as maximum power, maximum average power, and so on) can also be put on the signal vectors to analyze the communication system given in our paper. Since all symbols are assumed to be equally likely, we do not consider a system in which there can be unequal distribution of power among different channel vectors. In particular, some amplitude modulation systems (such as single-sideband AM) do not satisfy the requirement that channel vectors corresponding to different message characters have the same average power S, and hence such systems are not covered in this paper.

minate on the surface of a sphere of radius $(nS)^{\frac{1}{2}}$. By choosing the channel vectors appropriately, and by using maximum likelihood detection receiver, it can be shown[12] that the minimum probability of error $P(m, n, \rho^2)$, averaged over all symbols, satisfies the inequalities

$$Q(m, n, \rho^2) \leqq P(m, n, \rho^2) < \bar{Q}(m, n, \rho^2), \; \rho^2 = \frac{S}{N}, \tag{6}$$

where[8,12]

$$Q(m, n, \rho^2) = L\left(\theta_{m,n}, n, \rho^2 \frac{n}{2}\right), \tag{7}$$

$$\bar{Q}(m, n, \rho^2) = U\left(\theta_{m,n}, n, \rho^2 \frac{n}{2}\right), \tag{8}$$

$$\frac{2}{m} = \frac{\displaystyle\int_0^{\theta_{m,n}} \sin^{n-2} \mu \, d\mu}{\displaystyle\int_0^{\pi/2} \sin^{n-2} \mu \, d\mu} = I_{\sin^2 \theta_{m,n}}\left(\frac{n-1}{2}, \frac{1}{2}\right), \tag{9}$$

$$L\left(\theta, n, \rho^2 \frac{n}{2}\right) = \int_\theta^\pi p_n(\lambda) \, d\lambda, \tag{10}$$

$$U\left(\theta, n, \rho^2 \frac{n}{2}\right) = L\left(\theta, n, \rho^2 \frac{n}{2}\right) + \Omega(\theta) \int_0^\theta p_n(\lambda) \Omega(\lambda) \, d\lambda, \tag{11}$$

$$\Omega(\lambda) = \frac{(n-1)\pi^{(n-1)/2}}{\Gamma\left(\frac{n+1}{2}\right)} \int_0^\lambda \sin^{n-2} \mu \, d\mu, \tag{12}$$

$$\Gamma(k) = \int_0^\infty e^{-x} x^{k-1} \, dx, \tag{13}$$

$$p_n(\lambda) = \frac{(n-1) \exp\left(-\rho^2 \frac{n}{2} \sin^2 \lambda\right) \sin^{n-2} \lambda}{2^{n/2} (\pi)^{\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right)} \cdot \int_0^\infty r^{n-1} \exp\left[-\frac{[r - \rho(n)^{\frac{1}{2}} \cos \lambda]^2}{2}\right] dr, \tag{14}$$

and $I_x(\alpha, \beta)$ is the incomplete beta function given by

$$I_x(\alpha, \beta) = \frac{1}{B(\alpha, \beta)} \int_0^x t^{\alpha-1}(1-t)^{\beta-1} \, dt, \qquad 0 \leqq x \leqq 1; \tag{15}$$

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} \, dt. \tag{16}$$

The significance of the inequalities in equation (6) may be explained as follows. No m-ary digital modulation system with a given bandwidth expansion n and a given signal-to-noise ratio ρ^2 can achieve a lower probability of error than that given by $Q(m, n, \rho^2)$.* Also, we observe from equation (6) that m-ary digital modulation systems can be built to have an error probability given by $\bar{Q}(m, n, \rho^2)$.†

The bounds given by equation (6) can, therefore, be used in comparing the performance of practical modulation systems with that of the optimal systems, and in estimating the quality of performance of the practical modulation systems. For a given probability of error, and m and n, we can compare the signal-to-noise ratio required for a practical modulation system with the minimum signal-to-noise ratio predicted by equation (6) for the optimal system.‡

Since we are interested in this kind of comparison and since there seem to be theoretical considerations which show[10] that $\bar{Q}(m, n, \rho^2)$ is a very weak bound [it is shown in Ref. 10 that some explicit codes can be constructed to make $P(m, n, \rho^2)$ very close to $Q(m, n, \rho^2)$], we shall not discuss the upper bound $\bar{Q}(m, n, \rho^2)$ any more in this paper.

IV. EVALUATION OF LOWER BOUND $Q(m, n, \rho^2)$

For the sake of comparing the performance of proposed modulation systems with that of the optimal systems, it is essential to evaluate the lower bound $Q(m, n, \rho^2)$ for different values of m and n. The evaluation of $Q(m, n, \rho^2)$ is rather difficult and is usually done by using a digital computer. Slepian[8] has given methods of evaluating this bound when n is *odd*, and numerical values are given for $Q(m, n, \rho^2)$ when $1 \geqq Q(m, n, \rho^2) \geqq 10^{-5}$, $m = 32, 64, 128$ and 256, and $n \geqq 5$.

Since error rates of less than 10^{-5} are desired in digital systems and since no numerical results are available when $m < 32$ (in general, it is easier to build digital modulation systems with low values of m), we shall give some further numerical values for the cases considered by Slepian. In addition, we shall give a method to evaluate $Q(m, n, \rho^2)$ when n is even and point out its special significance when $n = 2$.

First we review briefly Slepian's method of evaluating $Q(m, n, \rho^2)$.[10] For any given values of m and n, we evaluate by interpolation the value of $\theta_{m,n} (0 < \theta_{m,n} \leqq \pi/2)$ from the set of tables for $I_x(\alpha, \beta)$ given in Ref.

* Of course, we assume that the digital modulation system satisfies other requirements given in this paper.

† For $n = 1$ or 2, it can be shown that we can make $Q(m, n, \rho^2) = P(m, n, \rho^2)$. Also, for all (integral) n, and $m = 2$, we can make $Q(2, n, \rho^2) = P(2, n, \rho^2)$.

‡ In making this comparison, we have to assume that we can estimate the bandwidth expansion factor n for the practical modulation system.

13. Since it has been shown[10] that

$$L(\theta, n, \sigma^2) = L(\theta, n-2, \sigma^2) + \cos\theta G(\theta, n-2, \sigma^2), \qquad n > 3, \tag{17}$$

$$\sigma^2 = \rho^2 \frac{n}{2}, \tag{18}$$

$$G(\theta, n, \sigma^2) = \sigma \cos\theta \sin\theta b_n G(\theta, n-1, \sigma^2) + \frac{n-2}{n-1} \sin^2\theta G(\theta, n-2, \sigma^2), \qquad n > 2, \tag{19}$$

$$b_n = \frac{n-2}{n-1} b_{n-2}, \qquad n > 2, \tag{20}$$

and

$$b_1 = \pi^{\frac{1}{2}}, \tag{21}$$

$$b_2 = \frac{2}{\pi^{\frac{1}{2}}}, \tag{22}$$

$$G(\theta, 1, \sigma^2) = \tfrac{1}{2} \exp(-\sigma^2 \sin^2\theta)[2 - \operatorname{erfc}(\sigma\cos\theta)], \tag{23}$$

$$G(\theta, 2, \sigma^2) = \frac{1}{\pi} \sin\theta \exp(-\sigma^2) + \frac{2\sigma}{\pi^{\frac{1}{2}}} \sin\theta \cos\theta G(\theta, 1, \sigma^2), \tag{24}$$

$$L(\theta, 3, \sigma^2) = \tfrac{1}{2} \operatorname{erfc}(\sigma) + \cos\theta G(\theta, 1, \sigma^2), \tag{25}$$

where

$$\operatorname{erfc}(x) = \frac{2}{\pi^{\frac{1}{2}}} \int_x^\infty \exp(-t^2)\, dt = 1 - \operatorname{erf}(x), \tag{26}$$

$Q(m, n, \rho^2)$ can be evaluated for *odd* n from equations (7), (9) and (17) through (25). However, for even n, we cannot use Slepian's method of evaluating $Q(m, n, \rho^2)$ unless we can find an explicit expression for $L(\theta, 2, \sigma^2)$.

Now it has been proved[14,15,16] that moderately high values of $n(2 < n < 5)$ are required for some digital modulation systems in order to optimize transmission rates per unit bandwidth. Since we would like to compare these systems (and other systems with similar bandwidth expansions) with the optimal systems, we shall first express $Q(m, n, \rho^2)$ explicitly for odd n, and $n \leqq 5$ before we discuss the evaluation of $Q(m, n, \rho^2)$ for even n.

4.1 *Lower Bound* $Q(m, n, \rho^2)$ *for* $n = 3, 5$

For $n = 3$, equations (7), (9), (10), (13) and (14) can be shown to yield

$$Q(m,\,3,\,\rho^2) = \frac{1}{2}\Bigg(\operatorname{erfc}\,[\rho(\tfrac{3}{2})^{\frac{1}{2}}] + (1 - 2/m)$$

$$\cdot \exp\left[-\frac{6\rho^2}{m}\left(1 - \frac{1}{m}\right)\right]\left\{2 - \operatorname{ercf}\left[\rho(\tfrac{3}{2})^{\frac{1}{2}}\left(1 - \frac{2}{m}\right)\right]\right\}\Bigg). \qquad (27)$$

For $m = 2^\ell$, $1 \leqq \ell \leqq 5$, we have evaluated $Q(m, 3, \rho^2)$ and have shown the results in Fig. 2.

Let us now compare the performance of an FSK system (using square-wave modulation, ideal discrimination detection with an integrate-and-dump circuit as tbe post-detection filter) with a bandwidtb expansion of 3 with the performance of the optimal system. For the FSK system the symbol error probability P_{FSK} can be shown[16,17] to be given by

$$P_{FSK} \sim \frac{1}{(2\pi\rho^2)^{\frac{1}{2}}}\,\frac{\cot\left\{\dfrac{\pi}{4}\,\dfrac{n-2}{m-1}\right\}}{\left\{\cos\left[\dfrac{\pi}{2}\,\dfrac{n-2}{m-1}\right]\right\}^{\frac{1}{2}}}\exp\left[-2\rho^2\sin^2\left\{\frac{\pi}{4}\,\frac{n-2}{m-1}\right\}\right],$$

$$\rho^2 \gg 1, \quad n < m + 1. \qquad (28)$$

For $n = 3$ and $m = 4$ and 8 we have plotted P_{FSK} and $Q(m, 3, \rho^2)$ in Fig. 3. Noting that the error rate of the optimal system can be made close to $Q(m, n, \rho^2)$ for small n, it can be observed from Fig. 3 that the error performance of the FSK system is inferior to the optimal system by several decibels. However, note that the formula in equation (28) is an asymptotic formula and that we have calculated the bandwidth expansion factor for the FSK system by Carson's rule. Also, note tbat we have taken the bandwidth of the message source to be $1/2T$, where T is the signaling interval of the source. If all tbese assumptions are reasonable, we must conclude from Fig. 3 that the performance of the FSK system is far from being optimum.

Let us now consider $n = 5$. For $n = 5$, we have

$$Q(m,\,5,\,\rho^2)$$

$$= \frac{1}{2}\Bigg[\operatorname{erfc}\,[\rho(\tfrac{5}{2})^{\frac{1}{2}}] + \frac{5\rho\sin^2\theta\cos^2\theta}{(10\pi)^{\frac{1}{2}}}\exp\{-\rho^2\tfrac{5}{2}\}$$

$$+ \tfrac{1}{2}\cos\theta\exp\{-\rho^2\tfrac{5}{2}\sin^2\theta\}(\sin^2\theta + 5\rho^2\cos^2\theta\sin^2\theta + 2)$$

$$\cdot\{2 - \operatorname{erfc}\,[\rho(\tfrac{5}{2})^{\frac{1}{2}}\cos\theta]\}\Bigg], \qquad (29)$$

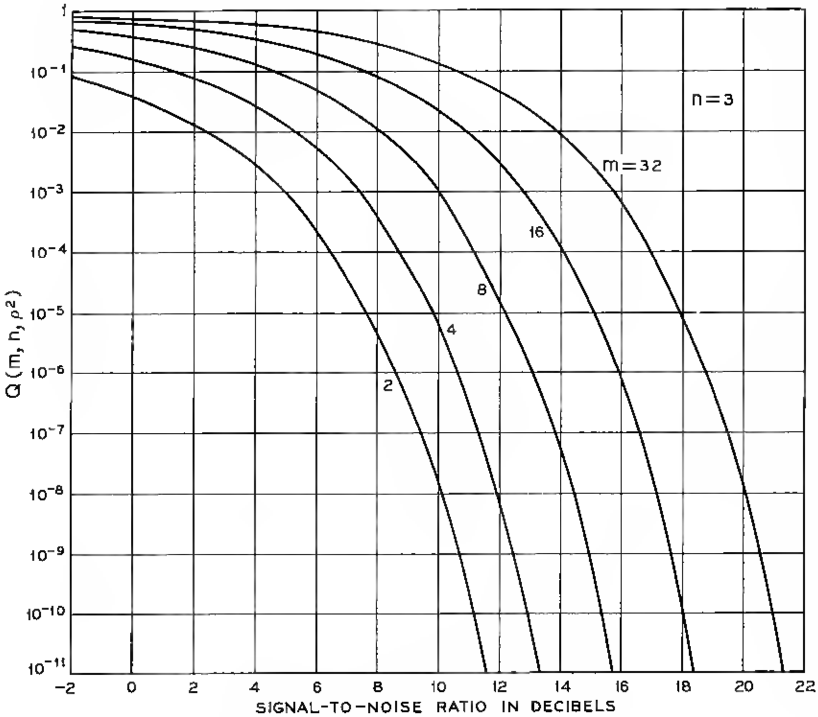

Fig. 2—Lower bound $Q(m, n, \rho^2)$ for $n = 3$.

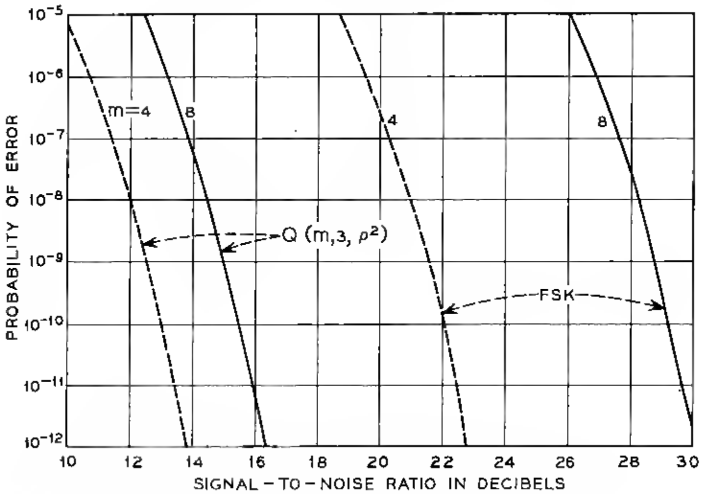

Fig. 3—Comparison of the FSK system with the optimal system, $n = 3$.

where

$$\cos\theta = 2\cos\left[\frac{4\pi}{3} + \frac{1}{3}\cos^{-1}\left(\frac{2}{m} - 1\right)\right], \tag{30}$$

and

$$\sin\theta = +(1 - \cos^2\theta)^{\frac{1}{2}}. \tag{31}$$

For $m = 2^\ell$, $1 \leqq \ell \leqq 5$, we have plotted $Q(m, 5, \rho^2)$ in Fig. 4.

For n odd, and $n > 5$, derivation of an expression for $Q(m, n, \rho^2)$ becomes rather tedious and long, and we shall not give these expressions. However, for $n = 7, 9, 11, 13$ and 17, we have calculated $Q(m, n, \rho^2)$ for $m = 2^\ell$, $1 \leqq \ell \leqq 5$, and the results are given in Figs. 5, 6, 7, 8 and 9. These numerical results which add to the results given by Slepian were obtained by using his method (see Appendix A for an alternative method).

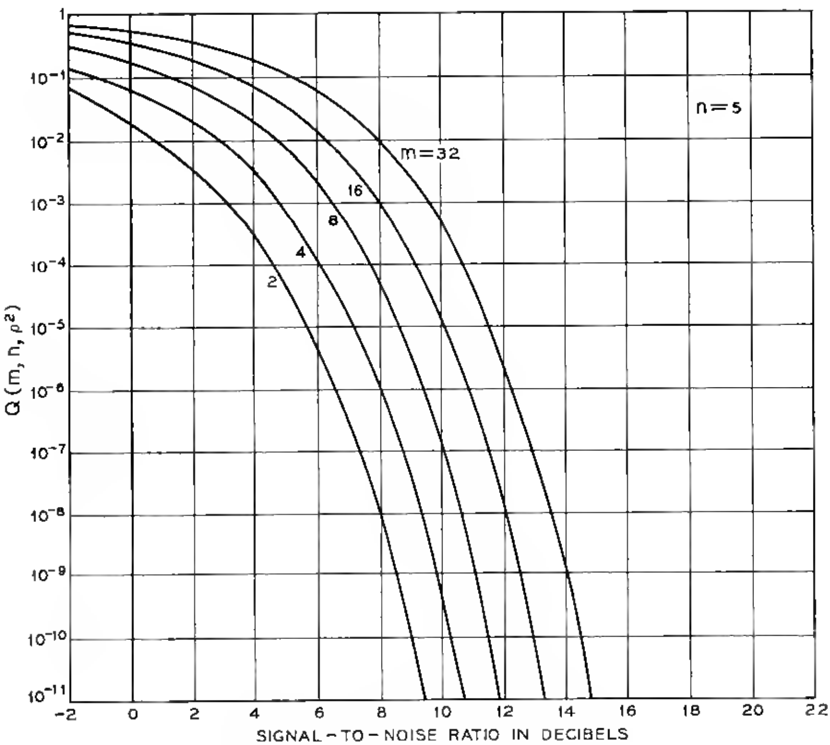

Fig. 4—Lower bound $Q(m, n, \rho^2)$ for $n = 5$.

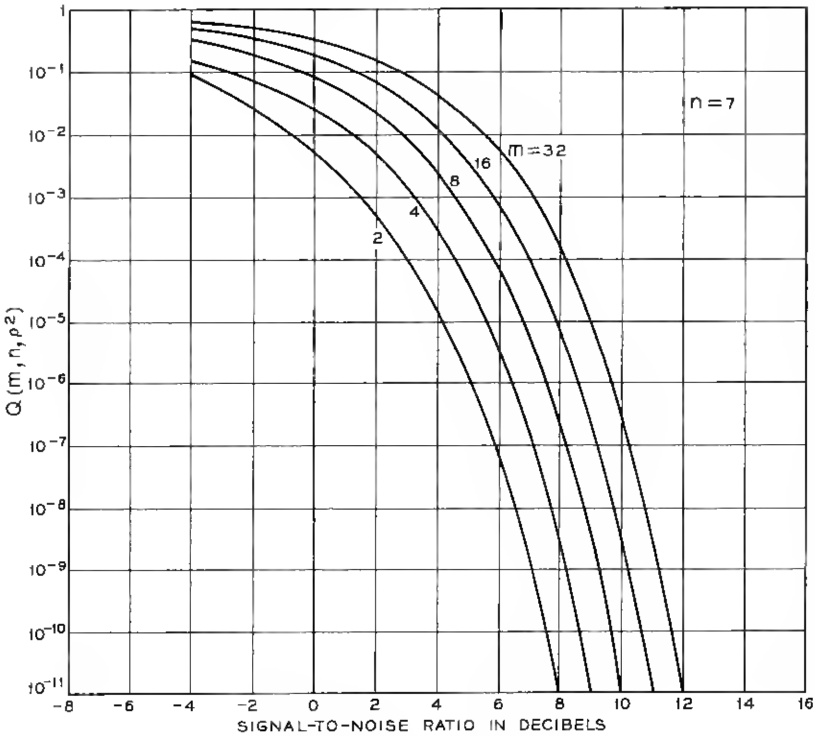

Fig. 5—Lower bound $Q(m, n, \rho^2)$ for $n = 7$.

4.2 *Lower Bound $Q(m, n, \rho^2)$ for n even*

Since we can calculate $Q(m, n, \rho^2)$ from $L(\theta, n, \sigma^2)$ and since the recurrence equation relates $L(\theta, n, \sigma^2)$ to $L(\theta, n - 2, \sigma^2)$ [see equation (17)], we can calculate $Q(m, n, \rho^2)$ for even n if we can calculate $L(\theta, 2, \sigma^2)$.

It can be shown that

$$L(\theta, 2, \sigma^2) = \int_\theta^\pi p_2(\lambda, \sigma^2)\, d\lambda \tag{32}$$

where

$$p_2(\lambda, \sigma^2) = \frac{1}{\pi}\left[e^{-\sigma^2} + \sigma(\pi)^{\frac{1}{2}} \cos \lambda e^{-\sigma^2 \sin^2 \lambda}\{1 + \operatorname{erf}(\sigma \cos \lambda)\}\right]. \tag{33}$$

Noting that $p_2(\lambda, \sigma^2)/2$ is the probability density of the phase angle of a sinusoidal carrier of amplitude $(2A)^{\frac{1}{2}}$ corrupted by random gaussian noise of average power N (signal-to-noise ratio $\sigma^2 = A/N$), we have

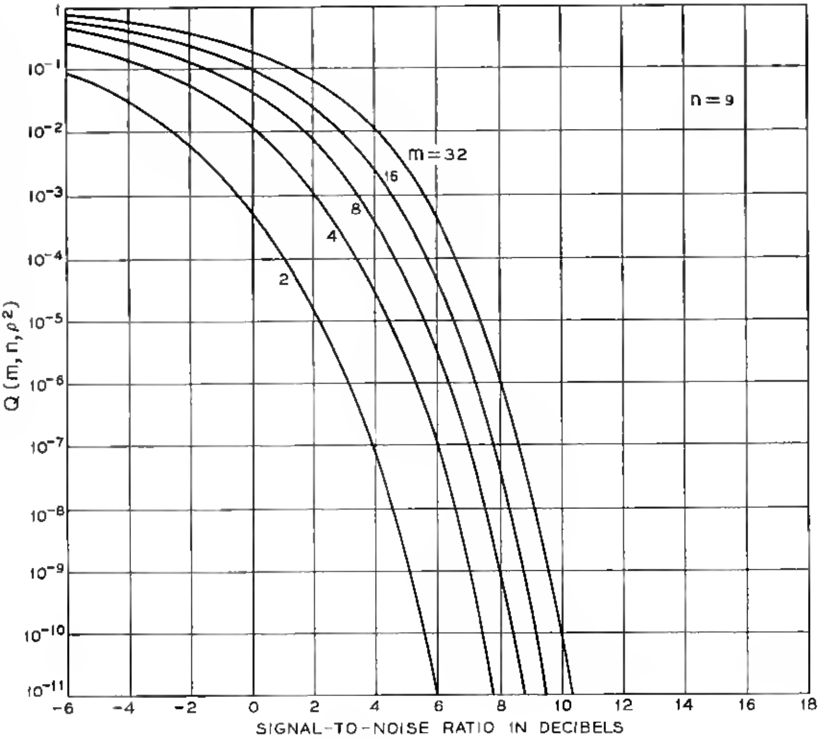

Fig. 6—Lower bound $Q(m,\, n,\, \rho^2)$ for $n = 9$.

shown in Appendix B that $L(\theta,\, 2,\, \sigma^2)$ can be calculated for any θ. Hence we can calculate $Q(m,\, n,\, \rho^2)$ for any even n. Now it can be shown[4] (see Appendix C) that*

$$L(\theta,\, 2,\, \sigma^2) = \tfrac{1}{2} \operatorname{erfc}(\sigma), \qquad \theta = \pi/2; \tag{34}$$

$$L(\theta,\, 2,\, \sigma^2) = \operatorname{erfc}\,[\sigma/(2)^{\frac{1}{2}}] - \tfrac{1}{4} \operatorname{erfc}^2 [\sigma/(2)^{\frac{1}{2}}], \qquad \theta = \pi/4$$

and

$$\begin{aligned} \tfrac{1}{2} \operatorname{erfc}(\sigma \sin \theta) + \max \Big\{ 0,\, \tfrac{1}{2} &\operatorname{erfc}(\sigma \sin \theta) \\ &- \frac{\tan \theta}{\pi} \exp(-\sigma^2)[1 - \pi^{\frac{1}{2}} \sigma \exp(\sigma^2) \operatorname{erfc}(\sigma)] \Big\} \\ &\leqq L(\theta,\, 2,\, \sigma^2) < \operatorname{erfc}(\sigma \sin \theta), \qquad 0 < \theta \leqq \pi/2; \end{aligned} \tag{35}$$

* Some of these results can be obtained from Ref. 4 by putting $\Omega = 0$.

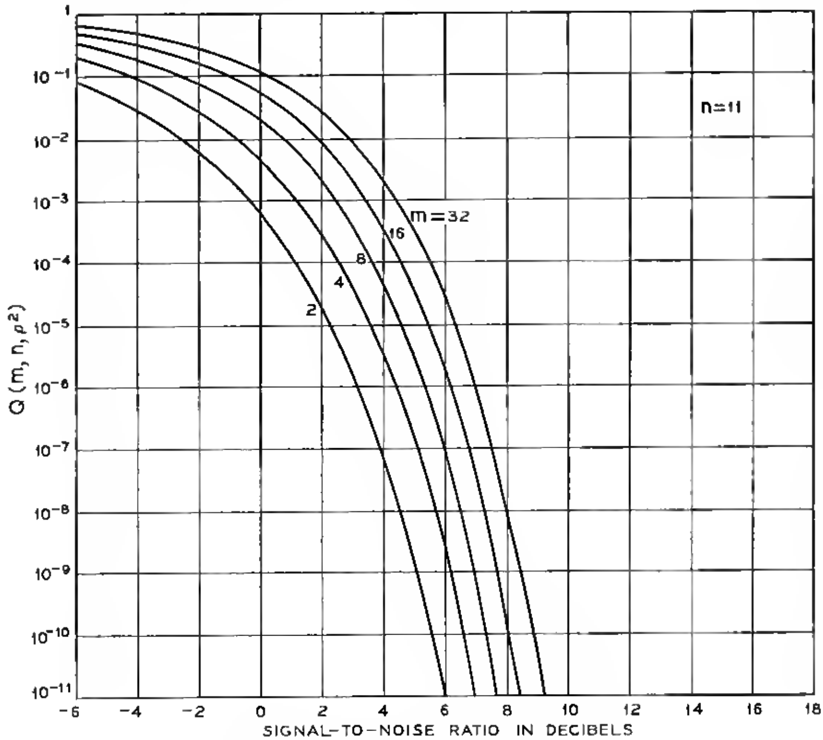

Fig. 7—Lower bound $Q(m, n, \rho^2)$ for $n = 11$.

where

$$\max \{a, b\} = \begin{cases} a, & a \geqq b; \\ b, & a < b. \end{cases} \tag{36}$$

Since the upper and lower bounds to $L(\theta, 2, \sigma^2)$ cannot differ by more than a factor of two and since all quantities involved in equations (17) through (25) are positive, we shall now write a modified bound

$$Q'(m, n, \rho^2) = L'(\theta_{m,n}, n, \sigma^2) \tag{37}$$

where

$$L'(\theta, n, \sigma^2) = L'(\theta, n-2, \sigma^2) + \cos\theta G(\theta, n-2, \sigma^2), \qquad n > 3, \tag{38}$$

and

$$L'(\theta, 2, \sigma^2) = \tfrac{1}{2}\,\mathrm{erfc}\,(\sigma \sin \theta). \tag{39}$$

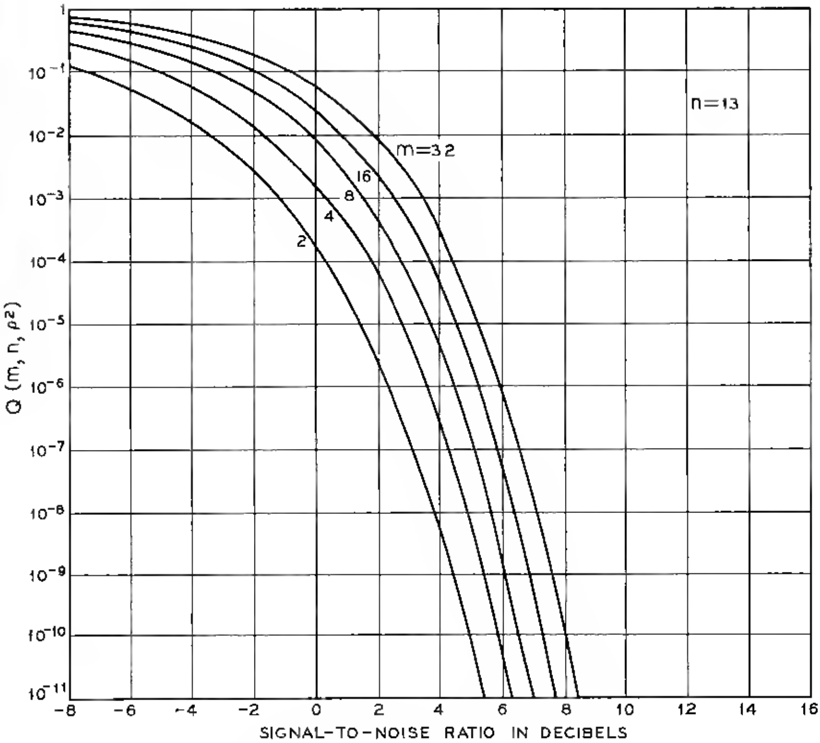

Fig. 8—Lower bound $Q(m, n, \rho^2)$ for $n = 13$.

Since

$$L(\theta, 2, \sigma^2) \geqq L'(\theta, 2, \sigma^2), \qquad 0 < \theta \leqq \pi/2 \tag{40}$$

note that

$$P(m, n, \rho^2) \geqq Q'(m, n, \rho^2). \tag{41}$$

Let us now consider the particular case $n = 2$. For $n = 2$,

$$\frac{2}{m} = \frac{\int_0^{\theta_{m,2}} d\mu}{\int_0^{\pi/2} d\mu} = \frac{2\theta_{m,2}}{\pi} \tag{42}$$

or

$$\theta_{m,2} = \frac{\pi}{m}. \tag{43}$$

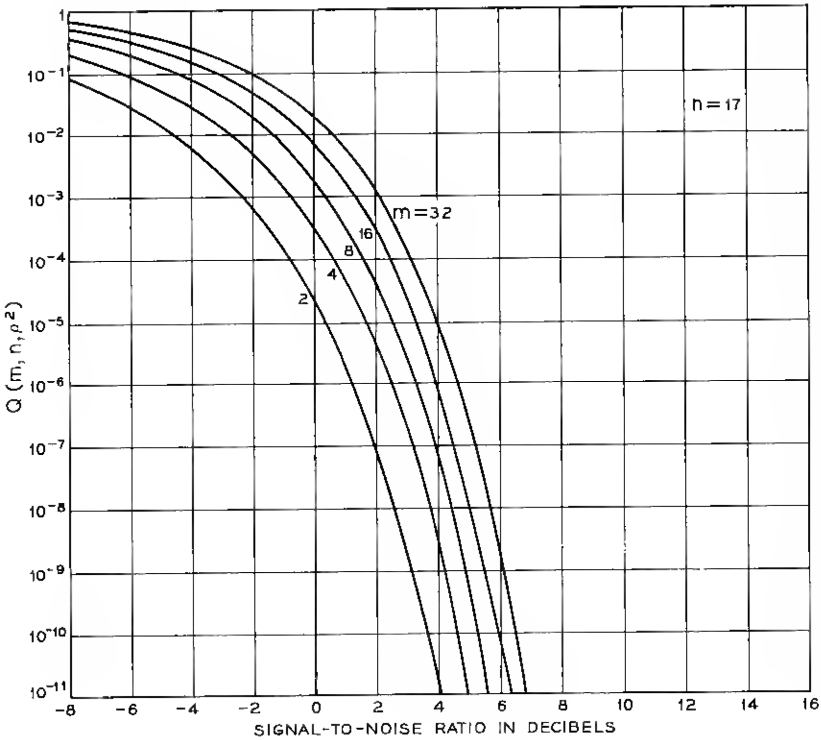

Fig. 9—Lower bound $Q(m, n, \rho^2)$ for $n = 17$.

Since $\theta_{m,2} = \pi/m$, and $p_2(\lambda, \sigma^2)$ is given by equation (33), it can be shown[18,19,20] from equations (7) and (10) that $Q(m, 2, \rho^2)$ is equal to the error probability obtained in m-ary coherent phase-shift keyed (CPSK) systems. Also, for $n = 2$, it can be shown[12] that

$$P(m, 2, \rho^2) = Q(m, 2, \rho^2). \tag{44}$$

It, therefore, follows that the error rates obtained in m-ary coherent phase-shift keyed systems are identical to those obtained in any m-ary digital modulation system that has a bandwidth expansion of two.* Hence we conclude that the error rates of any digital modulation system with a bandwidth expansion of two cannot be lower than the error rates of CPSK systems for all m.

Since the error rates of CPSK systems have been investigated in

* CPSK systems can be shown [20,21] to have approximately a bandwidth expansion of two.

detail,[18,19,20] we shall not give numerical values of $Q(m, 2, \rho^2)$ in this paper. However, we would like to note that[18,19,20]

$$Q(2, 2, \rho^2) = \tfrac{1}{2} \operatorname{erfc}(\rho), \tag{45}$$

$$Q(4, 2, \rho^2) = \operatorname{erfc}[\rho/(2)^{\frac{1}{2}}] - \tfrac{1}{4} \operatorname{erfc}^2 [\rho/(2)^{\frac{1}{2}}], \tag{46}$$

and

$$\begin{aligned}\tfrac{1}{2} \operatorname{erfc}(\rho \sin \pi/m) + \max \Big\{0, \tfrac{1}{2} \operatorname{erfc}(\rho \sin \pi/m) \\ - \frac{\tan \pi/m}{\pi} \exp(-\rho^2)[1 - \pi^{\frac{1}{2}}\rho \exp(\rho^2) \operatorname{erfc}(\rho)]\Big\} \\ \leqq Q(m, 2, \rho^2) < \operatorname{erfc}(\rho \sin \pi/m), \qquad m > 2.\end{aligned} \tag{47}$$

For signal-to-noise ratios greater than 5 dB, it can be shown[4] that

$$Q(m, 2, \rho^2) \approx \operatorname{erfc}(\rho \sin \pi/m), \qquad m \geqq 4, \tag{48}$$

and that the error in this approximation is less than 5 percent.

For $n > 2$, $Q(m, n, \rho^2)$ can be evaluated by using methods presented in Appendix B and using equations (17) through (25). However, this is usually difficult and tedious, and since $Q(m, n, \rho^2)$ and $Q'(m, n, \rho^2)$ can at most differ by a factor of two, we shall use the modified bound $Q'(m, n, \rho^2)$. Observe that $Q'(m, n, \rho^2)$ can easily be evaluated from equations (18) through (24), and (37) through (39).*

For $n = 4, 8, 12$ and 16, and $m = 2^\ell$, $1 \leqq \ell \leqq 5$, we have evaluated $Q'(m, n, \rho^2)$ and the results are shown in Figs. 10, 11, 12 and 13.

V. DISCUSSION AND CONCLUSIONS

Based on the results of Shannon and Slepian, we have derived, for different probabilities of error, lower bounds to the channel signal-to-noise ratio required by optimal systems to transmit the output of an m-ary message source through a channel of bandwidth expansion n. We assume that the channel is perturbed by additive white gaussian noise, all channel signals have the same average power S, and that the processing interval for decoding one message symbol is not greater than one signaling interval. When this interval can be arbitrary and when the transmission rate is not greater than the channel capacity, it is well known that the probability of error can be made arbitrarily

* For any n, note that $\theta_{2,n} = \pi/2$, and that $Q(2, n, \rho^2) = Q'(2, n, \rho^2) = \frac{1}{2} \operatorname{erfc} [\rho(n/2)^{1/2}]$.

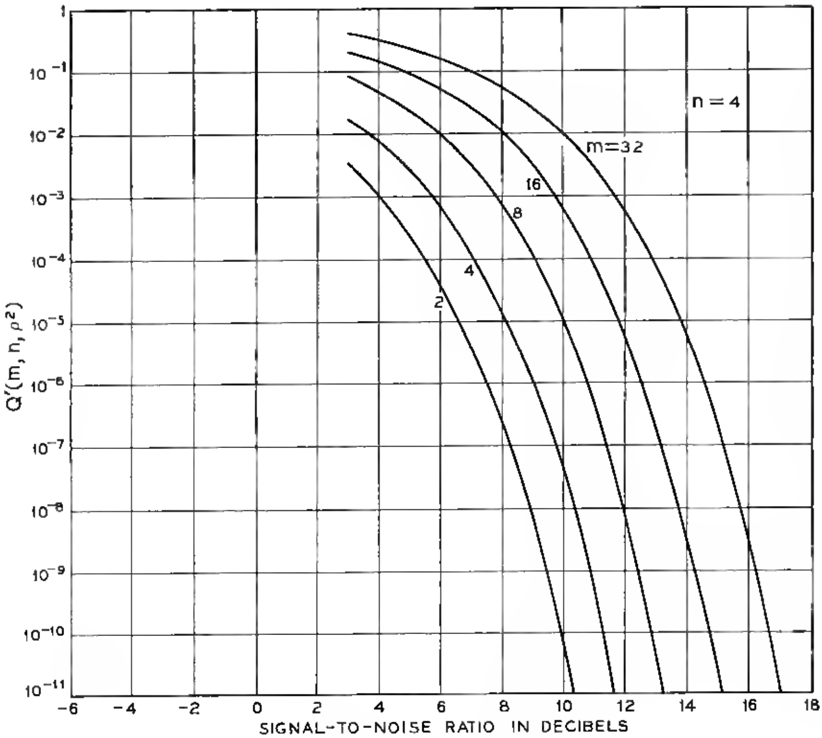

Fig. 10—Lower bound $Q'(m, n, \rho^2)$ for $n = 4$.

close to zero by using long and complex encoding and decoding procedures.

For different practical modulation systems, we can then compare the signal-to-noise ratio required for different probabilities of error with tbe lower bound given in this paper for optimal systems. This will aid us in deciding about the optimality or nonoptimality of different systems, and in evaluating the quality of performance of different modulation systems.

By using Slepian's method, we evaluate this lower bound for odd n, and $m = 2^{\ell}$, $1 \leqq \ell \leqq 5$. We also give a method of evaluating the lower bound for even n, and derive a simpler modified lower bound for n even, and $n > 2$. This modified lower bound has been evaluated for n even, and $m = 2^{\ell}$, $1 \leqq \ell \leqq 5$.

For a bandwidtb expansion of two, the performance of a coberent

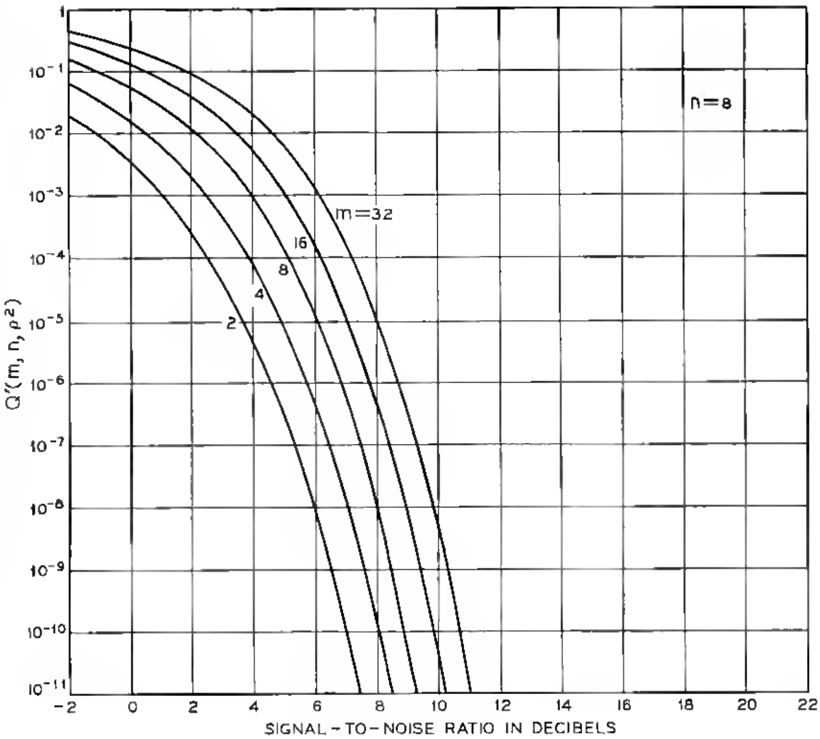

Fig. 11—Lower bound $Q'(m, n, \rho^2)$ for $n = 8$.

phase-shift keyed system has been shown to be as good as that of the optimal system.

A particular FSK system with a bandwidth expansion of three has been compared to the optimal system, and it appears that its performance is substantially inferior to that of the optimal system.

APPENDIX A

Evaluation of Lower Bound $Q(m, n, \rho^2)$

In this appendix, we shall give a second method to evaluate $Q(m, n, \rho^2)$. It can easily be shown from equation (9) that

$$\theta_{2,n} = \frac{\pi}{2} \quad \text{for all} \quad n, \tag{49}$$

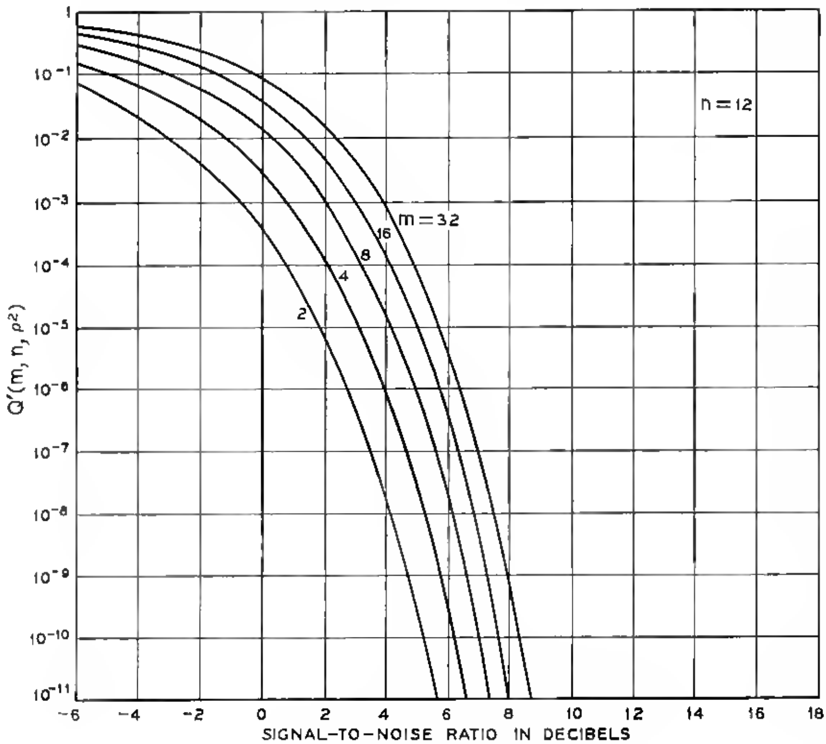

Fig. 12—Lower bound $Q'(m, n, \rho^2)$ for $n = 12$.

and that

$$P(2, n, \rho^2) = Q(2, n, \rho^2) = \tfrac{1}{2} \operatorname{erfc} [\rho(n/2)^{\frac{1}{2}}]. \tag{50}$$

Hence, we can write

$$L\left(\frac{\pi}{2}, n, \sigma^2\right) = \tfrac{1}{2} \operatorname{erfc} (\sigma). \tag{51}$$

Since

$$L(\theta, n, \sigma^2) = \int_\theta^\pi p_n(\lambda)\, d\lambda, \tag{52}$$

or

$$L\left(\theta, n, \rho^2 \frac{n}{2}\right) = \int_\theta^{\pi/2} p_n(\lambda)\, d\lambda + \int_{\pi/2}^{\pi} p_n(\lambda)\, d\lambda, \tag{53}$$

$$L\left(\theta, n, \rho^2 \frac{n}{2}\right) = T\left(\theta, n, \rho^2 \frac{n}{2}\right) + \tfrac{1}{2} \operatorname{erfc} [\rho(n/2)^{\frac{1}{2}}], \tag{54}$$

$$T(\theta, n, \sigma^2) = \frac{(n-1) \exp(-\sigma^2)}{2^{n/2}(\pi)^{\frac{1}{2}}\Gamma\left(\frac{n+1}{2}\right)} \int_0^\infty dr \int_\theta^{\pi/2} r^{n-1} \exp(-r^2/2)$$
$$\cdot \exp(r\sigma(2)^{\frac{1}{2}} \cos \lambda) \sin^{n-2} \lambda \, d\lambda. \tag{55}$$

Expanding $\exp(r\sigma(2)^{\frac{1}{2}} \cos \lambda)$ into a Taylor series and integrating term by term, we have

$$T(\theta, n, \sigma^2) = \frac{(n-1) \exp(-\sigma^2)}{2^{n/2}(\pi)^{\frac{1}{2}}\Gamma\left(\frac{n+1}{2}\right)} \cdot \sum_{\ell=1}^{\infty} \frac{[\sigma(2)^{\frac{1}{2}}]^\ell}{\ell!} \int_0^\infty r^{\ell+n-1} \exp - r^2/2_{dr}$$
$$\cdot \int_\theta^{\pi/2} \cos^\ell \lambda \sin^{n-2} \lambda \, d\lambda,$$
$$= \frac{(n-1) \exp(-\sigma^2)}{2(\pi)^{\frac{1}{2}}\Gamma\left(\frac{n+1}{2}\right)} \sum_{\ell=0}^{\infty} \frac{(2\sigma)^\ell}{\ell!} \Gamma\left(\frac{\ell+n}{2}\right)$$
$$\cdot \frac{1}{2} B\left(\frac{\ell+1}{2}, \frac{n-1}{2}\right) I_{\cos^2\theta}\left(\frac{\ell+1}{2}, \frac{n-1}{2}\right). \tag{56}$$

Equation (56) can be simplified to

$$T(\theta, n, \sigma^2) = \frac{\exp(-\sigma^2)}{2(\pi)^{\frac{1}{2}}} \sum_{\ell=0}^{\infty} \frac{(2\sigma)^\ell}{\ell!} \Gamma\left(\frac{\ell+1}{2}\right) I_{\cos^2\theta}\left(\frac{\ell+1}{2}, \frac{n-1}{2}\right). \tag{57}$$

Equations (54) and (57) yield

$$L(\theta, n, \sigma^2) = \tfrac{1}{2} \operatorname{erfc}(\sigma) + \frac{1}{2(\pi)^{\frac{1}{2}}} \exp(-\sigma^2)$$
$$\cdot \sum_{\ell=0}^{\infty} \frac{(2\sigma)^\ell}{\ell!} \Gamma\left(\frac{\ell+1}{2}\right) I_{\cos^2\theta}\left(\frac{\ell+1}{2}, \frac{n-1}{2}\right). \tag{58}$$

For m not too large and for large n, it can be shown that

$$\delta = \frac{\frac{\pi}{2} - \theta_{m,n}}{\pi/2} \tag{59}$$

is small. If δ is small, we can prove that the series given in equation

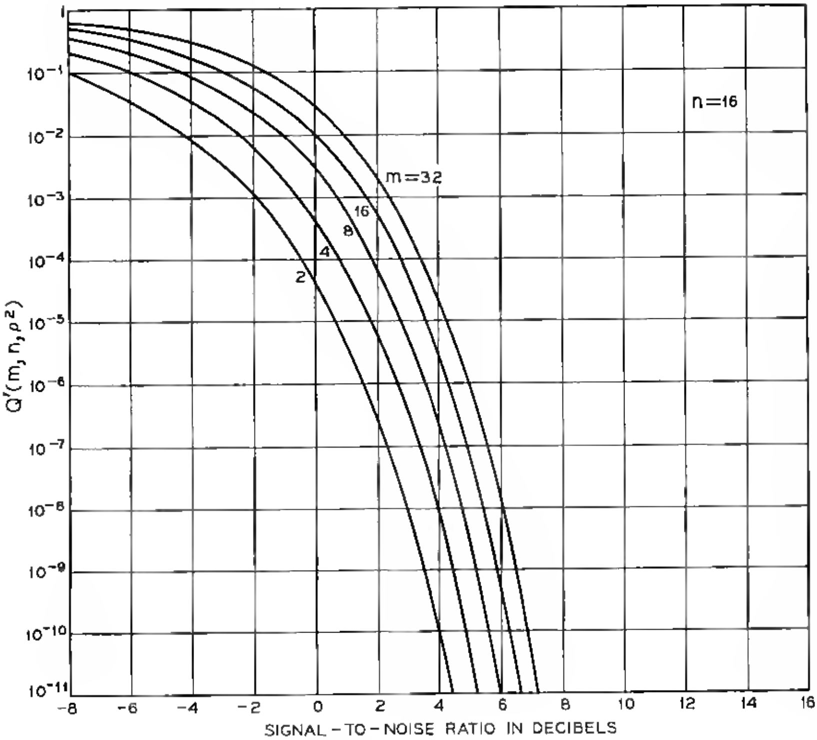

Fig. 13—Lower bound $Q'(m, n, \rho^2)$ for $n = 16$.

(58) converges rapidly and that we may alternatively calculate $Q(m, n, \rho^2)$ from equations (7), (9) and (58).

APPENDIX B

Evaluation of Distribution Function $L(\theta, 2, \sigma^2)$

Equations (10) and (14) can be shown to yield

$$L(\theta, 2, \sigma^2) = 1 - \int_{-\theta}^{\theta} p(\mu)\, d\mu, \tag{60}$$

where

$$p(\lambda) = \frac{1}{2\pi} [\exp(-\sigma^2) + \sigma(\pi)^{\frac{1}{2}} \cos \lambda \exp(-\sigma^2 \sin^2 \lambda) \cdot \{1 + \operatorname{erf}(\sigma \cos \lambda)\}]. \tag{61}$$

Note[18–20] that $p(\lambda)$ is the probability density function of the phase angle λ, $0 \leqq \lambda < 2\pi$, of the sum of a sinusoidal carrier (of unit amplitude) and gaussian noise [of average power $1/(2\sigma^2)$].

We can also show[19] that

$$L(\theta, 2, \sigma^2) = 1 - \frac{\theta}{\pi} - \sum_{k=1}^{\infty} \frac{2h_k}{k} \sin k\theta, \tag{62}$$

where

$$h_{2\ell+1} = \frac{(\pi\sigma^2)^{\frac{1}{2}}}{2\pi} \exp(-\sigma^2/2)[I_\ell(\sigma^2/2) + I_{\ell+1}(\sigma^2/2)], \qquad \ell = 0, 1, 2, \cdots \tag{63}$$

$$h_{2\ell} = \frac{1}{\pi} \sum_{n=-\infty}^{\infty} A_{2n+1} B_{2\ell-(2n+1)}, \qquad \ell = 1, 2, 3, \cdots ; \tag{64}$$

$$A_{-2s-1} = A_{2s+1} = \frac{(\pi\sigma^2)^{\frac{1}{2}}}{2} \exp(-\sigma^2/2)[I_s(\sigma^2/2) + I_{s+1}(\sigma^2/2)], \qquad s = 0, 1, 2, \cdots ; \tag{65}$$

and

$$B_{-2p-1} = B_{2p+1} = (-1)^p \frac{1}{\pi} \frac{(\pi\sigma^2)^{\frac{1}{2}}}{2p+1} \exp(-\sigma^2/2) \cdot [I_p(\sigma^2/2) + I_{p+1}(\sigma^2/2)], \qquad p = 0, 1, 2, \cdots . \tag{66}$$

$I_n(x)$ is the modified Bessel function of the first kind and of order n.

Since all h_k's can be calculated using either a set of tables or a digital computer and since the series given in equation (62) converges, we can calculate $L(\theta, 2, \sigma^2)$ for all σ and θ.

APPENDIX C

Evaluation of Upper and Lower Bounds

From equations (10) and (14), and Appendix B, we observe that $L(\theta, 2, \sigma^2)$ is the probability that the phase angle λ, $0 \leqq \lambda < 2\pi$, of a sinusoidal carrier of zero initial phase and unit amplitude lies outside the range $-\theta \leqq \lambda \leqq \theta$ when it is corrupted by random white gaussian noise of average power $1/2\sigma^2$.

When $\theta = \pi/2$, we can show[19,20] that

$$L(\theta, 2, \sigma^2) = \tfrac{1}{2} \operatorname{erfc}(\sigma), \qquad \theta = \pi/2. \tag{67}$$

When $\theta = \pi/4$, we can also show[18,20] that

$$L(\theta, 2, \sigma^2) = \text{erfc}\,[\sigma/(2)^{\frac{1}{2}}] - \tfrac{1}{4}\,\text{erfc}^2\,[\sigma/(2)^{\frac{1}{2}}]. \tag{68}$$

When $0 \leqq \theta < \pi/2$, let the sinusoidal carrier be represented by phasor OS in Fig. 14. Let x_u and x_v represent the in-phase and quadrature components of white gaussian noise corrupting the sinusoidal carrier. The quantity $L(\theta, 2, \sigma^2)$ is, therefore, given by the probability that the terminus of the vector **OT** lies in areas marked 1, 2 and 3.

We, therefore, have[4]

$$\text{erfc}\,(\sigma \sin \theta) - \frac{\tan \theta}{\pi} \exp\,(-\sigma^2)[1 - \pi^{\frac{1}{2}}\sigma \exp\,(\sigma^2)\,\text{erfc}\,(\sigma)] \leqq L(\theta, 2, \sigma^2) < \text{erfc}\,(\sigma \sin \theta). \tag{69}$$

Also, since $L(\theta, 2, \sigma^2)$ is greater than the probability that the terminus of the vector **OT** lies in areas marked 1 and 2 (or 2 and 3), we can write

$$L(\theta, 2, \sigma^2) > \tfrac{1}{2}\,\text{erfc}\,(\sigma \sin \theta). \tag{70}$$

Combining equations (69) and (70), we get equation (35).

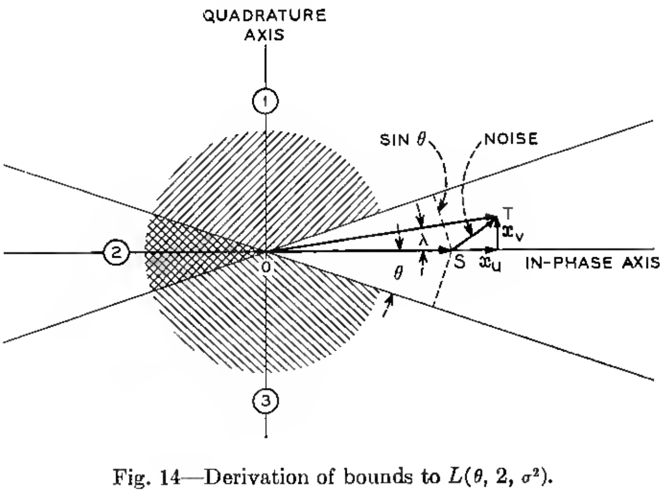

Fig. 14—Derivation of bounds to $L(\theta, 2, \sigma^2)$.

REFERENCES

1. Tillotson, L. C., "A Model of a Domestic Communication Satellite System," B.S.T.J., *47*, No. 10 (December 1968), pp. 2111–2136.
2. Tillotson, L. C., "Use of Frequencies Above 10 GHz for Common Carrier Applications," B.S.T.J., *48*, No. 6 (July–August 1969), pp. 1563–1576.
3. Ruthroff, C. L., and Tillotson, L. C., "Interference in a Dense Radio Network," B.S.T.J., *48*, No. 6 (July–August 1969), pp. 1727–1743.

4. Prabhu, V. K., "Error Rate Considerations for Coherent Phase-Shift Keyed Systems with Co-Channel Interference," B.S.T.J., *48*, No. 3 (March 1969), pp. 743–767.
5. Prabhu, V. K., and Enloe, L. H., "Interchannel Interference Considerations in Angle-Modulated Systems," B.S.T.J., *48*, No. 7 (September 1969), pp. 2333–2358.
6. Shannon, C. E., "Communication in the Presence of Noise," Proc. IRE, *37*, No. 1 (January 1949), pp. 1–12.
7. Shannon, C. E., "A Mathematical Theory of Communication," B.S.T.J., *27*, No. 3 (July 1948), pp. 379–423; *27*, No. 4 (October 1948), pp. 623–656.
8. Slepian, D., "The Threshold Effect in Modulation Systems that Expand Bandwidth," IRE Trans. on Information Theory, *IT-8*, No. 5 (September 1962), pp. 122–127.
9. Slepian, D., unpublished work.
10. Slepian, D., "Bounds on Communication," B.S.T.J., *42*, No. 3 (May 1963), pp. 681–707.
11. Landau, H. J., and Pollak, H. O., "Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty—III: The Dimension of the Space of Essentially Time- and Band-Limited Signals," B.S.T.J., *41*, No. 4 (July 1962), pp. 1295–1336.
12. Shannon, C. E., "Probability of Error for Optimal Codes in a Gaussian Channel," B.S.T.J., *38*, No. 3 (May 1959), pp. 611–656.
13. Pearson, K., *Tables of the Incomplete Beta-Function*, Cambridge University Press, London, England, 1934, pp. 1–15.
14. Pierce, J. R., unpublished work.
15. Rowe, H. E., unpublished work.
16. Mazo, J. E., Rowe, H. E., and Salz, J., "Rate Optimization for Digital Frequency Modulation," B.S.T.J., *48*, No. 9 (November 1969), pp. 3021–3030.
17. Mazo, J. E., and Salz, J., "Theory of Error Rates for Digital FM," B.S.T.J., *45*, No. 9 (November 1966), pp. 1511–1535.
18. Cahn, C. R., "Performance of Digital Phase-Modulation Communication Systems," IRE Trans. on Communication Systems, *CS-7*, No. 1 (May 1959), pp. 3–6.
19. Prabhu, V. K., "Error-Rate Considerations for Digital Phase-Modulation Systems," IEEE Trans. on Communication Technology, *COM-17*, No. 1 (February 1969), pp. 33–42.
20. Lucky, R. W., Salz, J., and Weldon, E. J., Jr., *Principles of Data Communication*, New York: McGraw-Hill, 1968, pp. 248–251.
21. Prabhu, V. K., unpublished work.